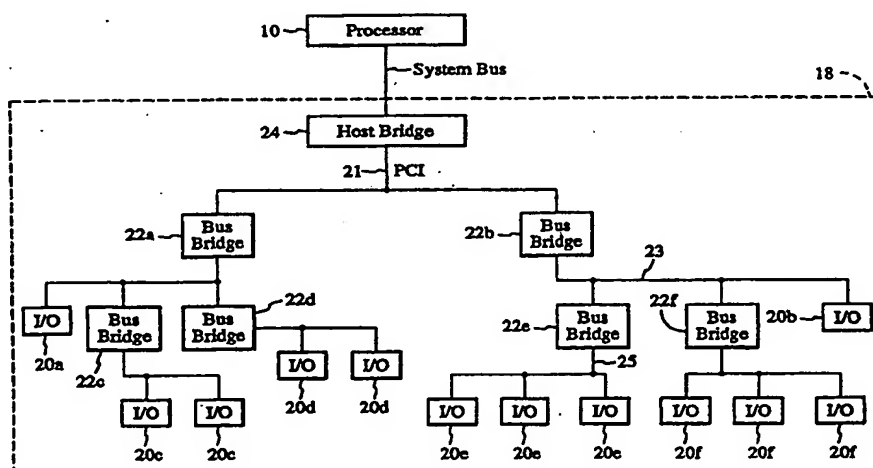




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G06F 11/22	A1	(11) International Publication Number: WO 98/44417 (43) International Publication Date: 8 October 1998 (08.10.98)
<p>(21) International Application Number: PCT/EP98/01674</p> <p>(22) International Filing Date: 23 March 1998 (23.03.98)</p> <p>(30) Priority Data: 08/829,088 31 March 1997 (31.03.97) US</p> <p>(71) Applicant (for AT BE CA CH CN CZ DE DK ES FI FR GB GR IE IT JP MC NL PL PT RU SE SG only): INTERNATIONAL BUSINESS MACHINES CORPORATION [US/US]; Old Orchard Road, Armonk, NY 10504 (US).</p> <p>(71) Applicant (for LU only): IBM DEUTSCHLAND INFORMATIONSSYSTEME GMBH [DE/DE]; D-70548 Stuttgart (DE).</p> <p>(72) Inventors: McLAUGHLIN, Charles, Andrew; 1203 Owl Court, Round Rock, TX 78681 (US). KITAMORN, Alongkorn; 16104 Fritsch Cove, Austin, TX 78717 (US).</p> <p>(74) Agent: TEUFEL, Fritz; IBM Deutschland Informationssysteme GmbH, Patentwesen und Urheberrecht, D-70548 Stuttgart (DE).</p>		<p>(81) Designated States: CA, CN, CZ, JP, PL, RU, SG, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p>Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>

(54) Title: A METHOD AND SYSTEM FOR FAULT ISOLATION FOR PCI BUS ERRORS



(57) Abstract

Method and system aspects for fault isolation on a bus are provided. In a method aspect, a method for isolating a fault condition on a bus of a computer system, the computer system including an input/output (I/O) subsystem formed by a plurality of I/O devices communicating via the bus, includes categorizing, in a recursive manner, the I/O subsystem, and isolating a source of an error condition within the I/O subsystem. Further, the I/O subsystem communicates via a peripheral component interconnect, PCI, bus. In a system aspect, a computer system for isolating a fault condition on a PCI bus includes a processing mechanism, and an input/output mechanism, coupled to the processing mechanism, comprising a plurality of input/output devices and bridges coupled to a PCI bus and communicating according to a PCI standard. In addition, the system includes a fault isolation mechanism within the processing mechanism for identifying a source of an error condition in the input/output mechanism. Further, the fault isolation mechanism performs categorization of the input/output mechanism in a recursive manner.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

D E S C R I P T I O N

A METHOD AND SYSTEM FOR FAULT ISOLATION FOR PCI BUS ERRORS

CROSS-REFERENCE TO RELATED APPLICATIONS

The present application is related to applications Serial No. 08/829,017, entitled "Method and System for Check Stop Error Handling," filed March 31, 1997; Serial No. 08/829,018, entitled "Error Collection Coordination for Software-Readable and Non-Software Readable Fault Isolation Registers in a Computer System," filed March 31, 1997; Serial No. 08/829,016, entitled "Machine Check Handling for Fault Isolation in a Computer System," filed March 31, 1997; Serial No. 08/829,089, entitled "Method and System for Reboot Recovery," filed March 31, 1997; and Serial No. 08/829,090, entitled "A Method and System for Surveillance of Computer System Operations," filed March 31, 1997.

FIELD OF THE INVENTION

The present invention relates generally to input/output operations in a computer system, and more particularly to fault isolation in a peripheral component interface (PCI) structure.

BACKGROUND OF THE INVENTION

In many computer systems, support of peripheral devices, such as hard disk drives, speakers, CD-ROM drives, etc., occurs through a standard I/O (input/output) device architecture called Peripheral Component Interconnect (PCI). The PCI architecture supports many complex features, including I/O expansion through PCI-to-PCI bridges, peer-to-peer (device-to-device) data transfers between controlling devices, i.e., masters, and responding devices, i.e., targets, as well as multi-function devices, and both integrated and plug-in devices.

The PCI architecture also defines standards for the detection and capture of error conditions on a PCI bus and in the devices. While the standard facilities provide error capture capabilities, the number of failure scenarios that may occur is large given the wide range of features allowed by the PCI architecture. Thus, isolating faults to a specific failing component becomes very difficult.

For example, for each transaction that occurs on the PCI bus, there is a master device which controls the transaction, and a target device which responds to the master's request. Since data can flow in either direction (i.e., the master can request a read or write), it is important to know which device was the sender of bad data and which device was the receiver. Also, since errors can flow across PCI-to-PCI bridges, it is important to know whether the fault is located on the near or far side of the bridge.

Accordingly, a need exists for a failure isolation technique that would operate successfully for the numerous options supported by the PCI architecture, while providing consistent diagnostic information to servicers across a wide variety of hardware platforms.

SUMMARY OF THE INVENTION

The present invention meets this need and provides method and system aspects for fault isolation on a PCI bus. In a method aspect, a method for isolating a fault condition on a bus of a computer system, the computer system including an input/output (I/O) subsystem formed by a plurality of I/O devices communicating via the bus, includes categorizing, in a recursive manner, the I/O subsystem, and isolating a source of an error condition within the I/O subsystem. Further, the I/O subsystem communicates via a peripheral component interconnect, PCI, bus.

In a further method aspect, a method for fault isolation for bus errors includes the steps of (a) processing a device error on a PCI bus, and (b) performing ordered categorization of a plurality of input/output devices coupled to the PCI bus. The method further includes (c) determining whether the device error originates from a subordinate branch of the PCI bus, and (d) recursively performing steps (a) - (c) until the PCI bus is categorized.

In a system aspect, a computer system for isolating a fault condition on a bus includes a processing mechanism, and an input/output mechanism coupled to the processing mechanism. The input/output mechanism comprises a plurality of input/output devices and bridges coupled to a PCI bus and communicating according to a PCI standard. In addition, the system includes a fault isolation mechanism within the processing mechanism for identifying a source of an error condition in the input/output mechanism. Further, the fault isolation mechanism performs categorization of the input/output mechanism in a recursive manner.

With the present invention, a fault isolation technique successfully provides more specific identification of an error source in a PCI bus architecture. The fault isolation technique greatly reduces the ambiguity of error occurrence when the numerous options supported by the PCI architecture are utilized in a given system. Further, by relying on the standard features of the PCI architecture, the fault isolation technique is readily applicable to varying system arrangements to provide versatile application. These and other advantages of the aspects of the present invention will be more fully understood in conjunction with the following detailed description and accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a block diagram of a computer system in accordance with the present invention.

Figure 2 illustrates a block diagram of an input/output subsystem of the computer system of Figure 1.

Figure 3 illustrates a flow diagram of a fault isolation process in accordance with the present invention.

Figure 4 illustrates a flow diagram of an ordered categorization step of Figure 2 in greater detail.

DESCRIPTION OF THE INVENTION

The present invention relates to fault isolation for a PCI architecture. The following description is presented to enable one of ordinary skill in the art to make and use the invention and is provided in the context of a patent application and its requirements. Various modifications to the preferred embodiment will be readily apparent to those skilled in the art and the generic principles herein may be applied to other embodiments. Thus, the present invention is not intended to be limited to the embodiment shown but is to be accorded the widest scope consistent with the principles and features described herein.

Figure 1 illustrates a basic block diagram of a general purpose computer system for use with the present invention. As shown, the computer system includes a processor 10, such as a PowerPC processor from IBM Corporation, Inc., coupled to memory 12, i.e., RAM (random access memory) and ROM (read only memory). An operating system (O/S) 14 typically runs on the processor to perform basic tasks in the computer system and act as a platform for application programs. Also included is firmware 16 that runs on the processor 10 and is code stored in suitable memory, such

as Flash memory, non-volatile RAM, or EPROM (erasably programmable read only memory), as is well understood to those skilled in the art. Further, an input/output (I/O) subsystem 18 is coupled to the processor 10 for controlling the interactions between the processor 10 and input/output devices, e.g., a hard disk drive, a monitor, etc., according to a PCI (peripheral component interface) standard.

Figure 2 presents an expanded illustration of the I/O subsystem 18 of the computer system of Figure 1. Of course, the number and types of components illustrated is meant to be illustrative and not restrictive of an embodiment of the present invention. Utilizing a PCI bus 21 allows a subsystem of I/O devices 20a-20f to interact with the processor 10. In utilizing a plurality of I/O devices 20a-20f, bridges 22a-22f support communication among the plurality of I/O devices 20a-20f with a host bridge 24 acting as a main link to the processor 10. Further, for the hierarchy of the I/O subsystem 18, primary buses and secondary buses exist for bridges linked with other bridges, e.g., primary bus 23 and secondary bus 25 for bridge 22b linked with bridge 22e. With the large number of bridges 22 and I/O devices 20 capable of co-existing in the computer system through the PCI architecture, the types and numbers of failure situations that can occur is high. With the present invention, isolation of a cause for a fault condition provides a significant improvement for diagnostic operations.

Figure 3 illustrates a general flow chart for failure isolation in accordance with the present invention. Preferably, the failure isolation is provided as a portion of the firmware 16 (Fig. 1), as is well appreciated by those skilled in the art. The process of isolating a fault condition suitably begins at a top-level PCI bus, i.e., the PCI bus directly under the host bridge 24 (Fig. 2), (step 30). The process continues (step 32) with an ordered categorization of the devices and components within the I/O subsystem. The ordered categorization relies on determining the status of the devices on the bus being examined

according to information available in architected status registers provided in the devices in accordance with PCI standards. Details of the ordered categorization are presented with reference to Figure 4. Generally, the ordered

5 categorization follows a specific order in a process of elimination manner to take into consideration all of the possibilities for errors that exist for data propagation within the hierarchical tree structure of the I/O subsystem 18.

10 Two major errors on PCI buses include PERR, parity error, which is signalled when a bad data parity condition is seen on the bus, and SERR, system error, which is signalled when an address parity error occurs or when a device has a critical error. Generating parity is non-optional, since it must be performed by

15 all PCI compliant devices. The target device for a particular PCI transaction checks parity and reports an address parity error. With respect to data parity errors, the master device detects and reports data parity errors for a particular read transaction, while the target device detects and reports data

20 parity errors for a particular write transaction. A master device, however, has the ability to detect an error whether the master or target device generated the error. Through the categorization of the present invention, isolation of both of these error conditions preferably occurs.

25 Referring to Figure 4, an examination for categorization (step 39) is made as to whether a PCI-to-PCI bridge received an SERR# signal on its secondary bus. Next an examination (step 40) for categorization occurs for a PCI-to-PCI bridge which received bad parity on its secondary bus. When that condition exists, a next examination (step 42) occurs for a PCI-to-PCI bridge which

30 received bad parity on its primary bus. The examination (step 44) continues with consideration for a PCI-to-PCI bridge acting as a master device to a target on the secondary bus which detected bad parity. A next examination (step 46) occurs for a

35 PCI-to-PCI bridge acting as a master device to a target on the primary bus which detected bad parity. Categorization continues

with an examination for a PCI-to-PCI bridge through which a target or master Abort was signalled (step 48).

Following examination of PCI-to-PCI bridges, categorization
5 continues with identification of a master device that detected bad parity (step 50). Further categorization occurs with identification of a master device of a target that detected bad parity (step 52). A next categorization examination occurs for a device that signalled SERR# due to bad address parity (step 54).
10 Subsequently, examination occurs for a master device that signalled SERR# due to a target Abort (step 56), and a master device that signalled SERR# due to a master Abort (step 58). Categorization continues by examining for a device that signalled SERR# due to an internal error (step 60), a target
15 device that detected bad parity (step 62), and a device that detected bad parity, but had SERR# reporting disabled (step 64). Additionally, categorization occurs with examination for a target device that signalled a target Abort (step 66), and for a potential sender of bad address parity, if other devices on the
20 bus are signalling detection of bad address parity (step 68).

As the categorization of a bus is occurring, the path of the error condition is followed. Referring back to Figure 3, when the categorization (step 32) indicates that a PCI-to-PCI bridge
25 connects to another PCI-to-PCI bridge from which the error condition is occurring (step 34), the sequence returns to perform the categorization on the bus supported by the other PCI-to-PCI bridge. Thus, the categorization is performed recursively from the top-level PCI bus down through all of the
30 sub-bridges, i.e., subordinate branches of the PCI bus through the hierarchy of the I/O subsystem 18 (Fig. 2). Once the ordered categorization is completed, the resulting information is preferably returned as an error log and analyzed for an
35 error/fault source isolation (step 36) within the I/O subsystem 18. With the similarity among error register values for many of the error conditions, the ordered categorization of the present

invention properly identifies the type of error each device may have detected to assist in the analysis of the fault source.

By way of example, address parity errors that result in an SERR# signal are isolated by finding a single device on a bus which did not detect bad address parity, since the only such device would be the one that issued the bad address on the bus. Data parity errors are isolated by finding the master and target devices, then determining which of the two actually detected the bad data. The device detecting the bad data is termed the "signalling" device, while the source of the bad data is termed the "sending" device. The "sending" device is the top priority for replacement. If the master and target are on two different buses (with one or more PCI-to-PCI bridges on the path between them), the failure is isolated to a specific bus. As a result, PCI-to-PCI bridges may be listed as the "sending" or "signalling" device, or both.

Further, for multi-function devices, examination suitably occurs as though there are distinct devices isolated to a same physical location. Further, internal device errors reported by an SERR# (system error, active low) signal are isolated to the signalling device. Additionally, aborted operations that result in an SERR# signal are suitably isolated to the master and target device, with the top priority for replacement being the device that caused the abort.

Although the present invention has been described in accordance with the embodiments shown, one of ordinary skill in the art will readily recognize that there could be variations to the embodiments and those variations would be within the spirit and scope of the present invention. By way of example, although the present invention is described in terms of a PCI bus, the fault isolation techniques are suitable for application with other bus structures, as well. Accordingly, many modifications may be made by one of ordinary skill in the art without departing from the spirit and scope of the appended claims.

CLAIMS

What is claimed is:

- 5 1. A method for isolating a fault condition on a bus of a computer system, the computer system including an input/output (I/O) subsystem formed by a plurality of I/O devices communicating via the bus, the method comprising the steps of:
 - 10 (a) categorizing, in a recursive manner, the I/O subsystem; and
 - (b) isolating a source of an error condition within the I/O subsystem.
- 15 2. The method of claim 1 wherein the I/O subsystem communicates via a peripheral component interconnect, PCI, bus.
- 20 3. The method of claim 2 wherein the I/O subsystem further comprises a PCI-to-PCI bridge, the PCI-to-PCI bridge having a primary bus and a secondary bus.
- 25 4. The method of claim 1 wherein categorizing step (a) further comprises examining whether a PCI-to-PCI bridge received a SERR# signal on the secondary bus.
- 30 5. The method for isolating of claim 4 wherein the categorizing step (a) further comprises examining for bad parity received on the secondary bus of the PCI-to-PCI bridge.
- 35 6. The method for isolating of claim 5 wherein the categorizing step (a) further comprises examining for bad parity received on the primary bus of the PCI-to-PCI bridge.

7. The method for isolating of claim 6 wherein the categorizing step (a) further comprises examining for the PCI-to-PCI bridge acting as a master device to a target device on the secondary bus which detected bad parity.
- 5
8. The method for isolating of claim 7 wherein the categorizing step (a) further comprises examining for the PCI-to-PCI bridge acting as the master device to the target device on the primary bus which detected bad parity.
- 10
9. The method for isolating of claim 8 wherein the categorizing step (a) further comprises examining for the PCI-to-PCI bridge signalling an abort.
- 15
10. The method for isolating of claim 9 wherein the categorizing step (a) further comprises examining for the master device detecting bad parity.
- 20
11. The method for isolating of claim 10 wherein the categorizing step (a) further comprises examining for the master device of the target device detecting bad parity.
- 25
12. The method for isolating of claim 11 wherein the categorizing step (a) further comprises examining for a device signalling a system error due to bad address parity.
- 30
13. The method for isolating of claim 12 wherein the categorizing step (a) further comprises examining for the master device signalling the system error due to an abort on the target device.
- 35
14. The method for isolating of claim 13 wherein the categorizing step (a) further comprises examining for the master device signalling the system error due to a master abort.

15. The method for isolating of claim 14 wherein the categorizing step (a) further comprises examining for the device signalling the system error due to an internal error.
- 5
16. The method for isolating of claim 15 wherein the categorizing step (a) further comprises examining for the target device detecting bad parity.
- 10
17. The method for isolating of claim 16 wherein the categorizing step (a) further comprises examining for a device detecting bad parity while system error reporting is disabled.
- 15
18. The method for isolating of claim 17 wherein the categorizing step (a) further comprises examining for the target device signalling a target abort.
- 20
19. The method for isolating of claim 18 wherein the categorizing step (a) further comprises examining for a potential sender of bad address parity.
- 25
20. A computer system for isolating a fault condition on a peripheral component interconnect, PCI, bus, the system comprising:
- a processing means;
- an input/output means coupled to the processing means and comprising a plurality of input/output devices and bridges coupled to a PCI bus and communicating according to a PCI standard; and
- 35
- fault isolation means within the processing means for identifying a source of an error condition in the input/output means.

21. The system of claim 20 wherein the fault isolation means further performs categorization of the input/output means in a recursive manner.

5 22. The system of claim 21 wherein the fault isolation means further provides an error log for isolation of the source of the error condition within the input/output means.

10 23. The system of claim 21 wherein the fault isolation means performs categorization by examining error condition values.

15 24. The system of claim 23 wherein the error condition values are stored in status registers of the input/output means.

25 25. A method for fault isolation for peripheral component interconnect (PCI) bus errors, the method comprising the steps of:

20 (a) processing a device error on a PCI bus;

(b) performing ordered categorization of a plurality of input/output devices coupled to the PCI bus;

25 (c) determining whether the device error originates from a subordinate branch of the PCI bus; and

(d) recursively performing steps (a) - (c) until the PCI bus is categorized.

30 26. The method of claim 25 further comprising forming an error log from the ordered categorization.

35 27. The method of claim 26 further comprising analyzing the error log to isolate the device error.

28. The method of claim 25 wherein the ordered categorizing examines status registers of the plurality of input/output devices.
- 5 29. The method of claim 28 wherein the plurality of input/output devices comprise one or more PCI-to-PCI bridge devices.
- 10 30. The method of claim 29 wherein the one or more PCI-to-PCI bridge devices support one or more subordinate branches of the PCI bus.

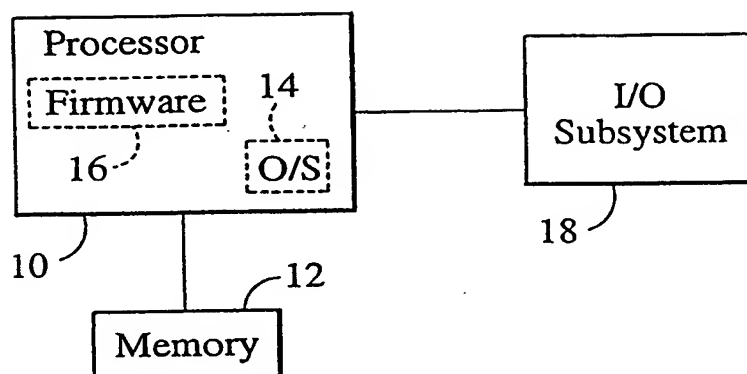


FIG. 1

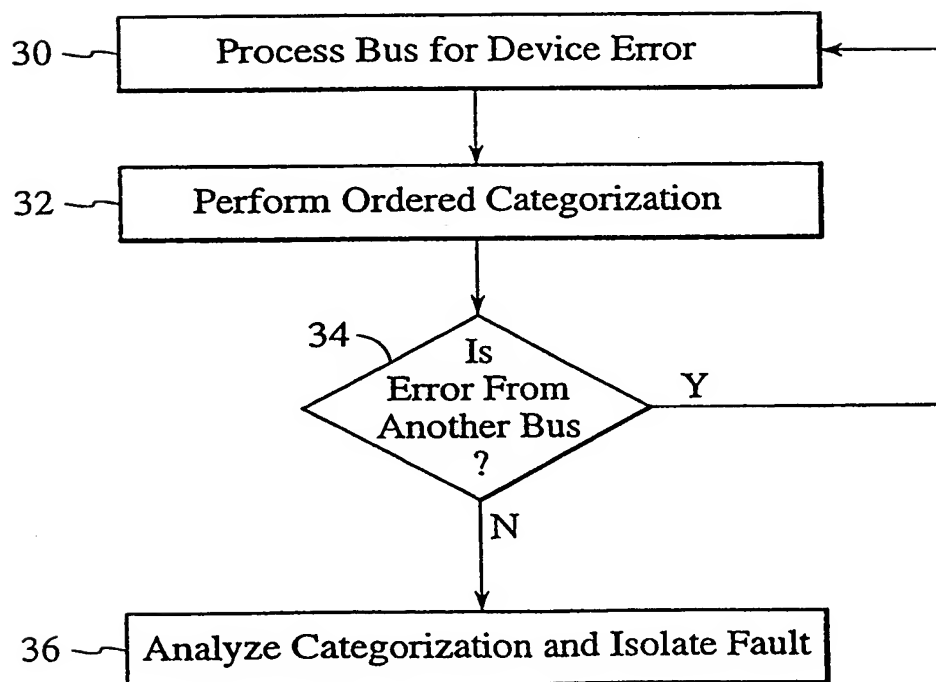


FIG. 3

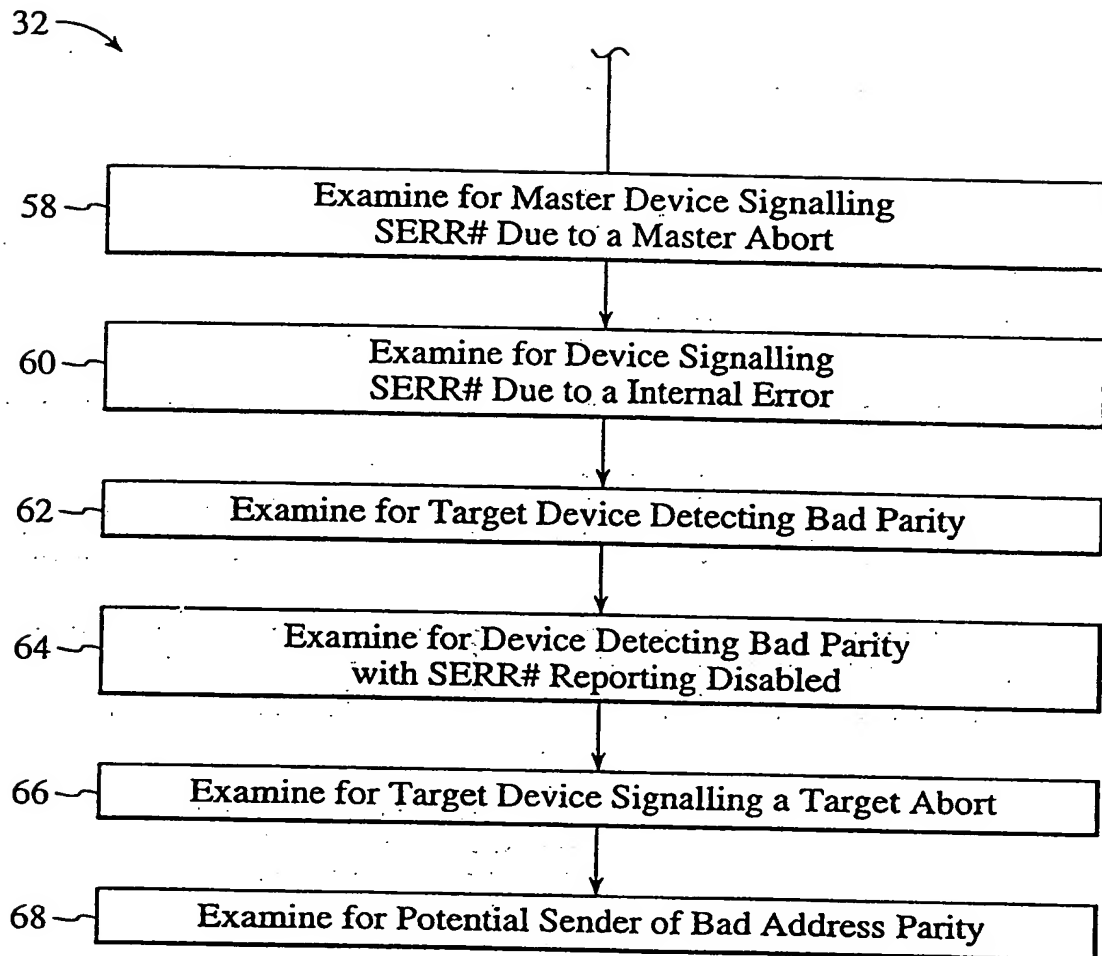
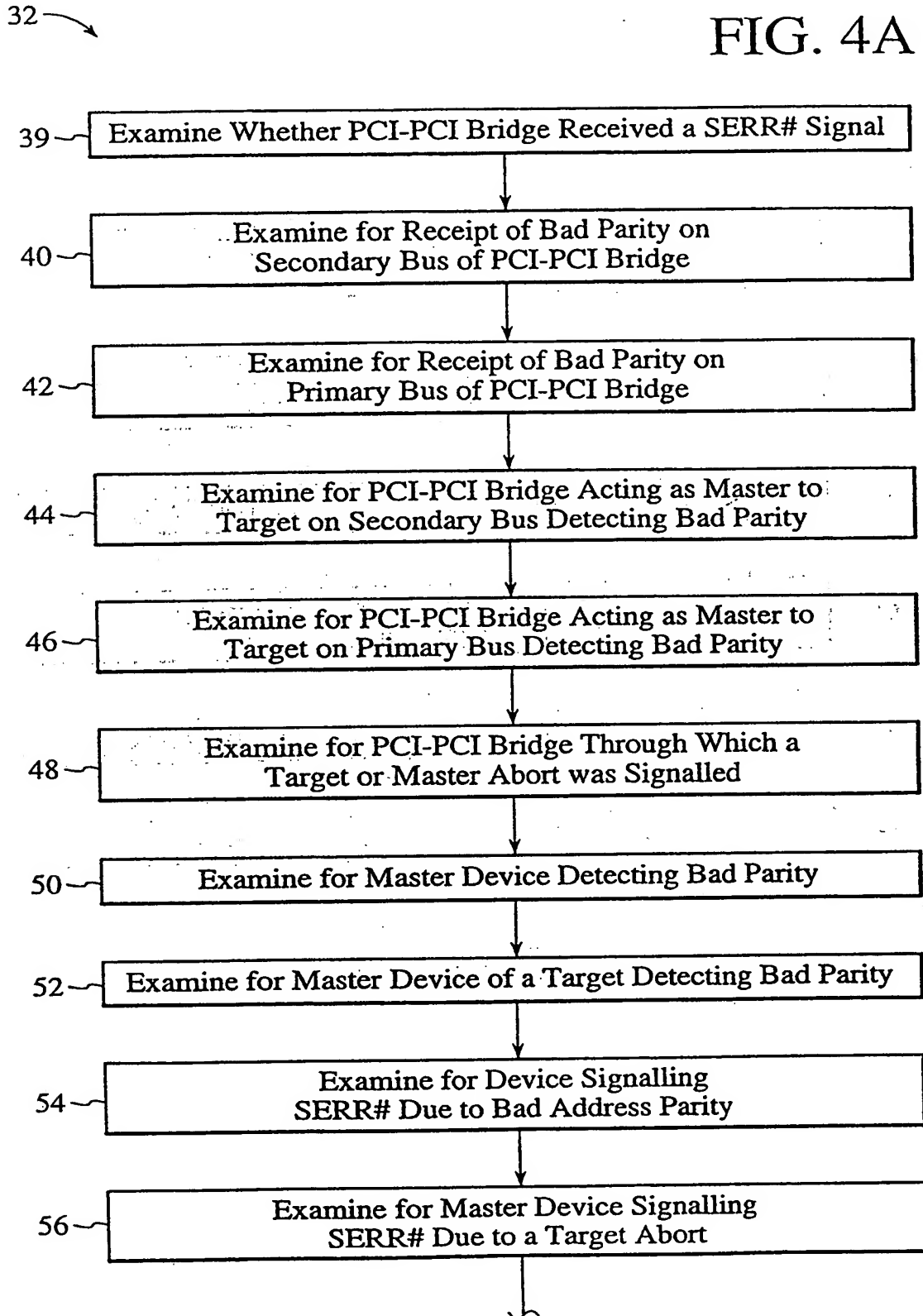


FIG. 4B

FIG. 4A



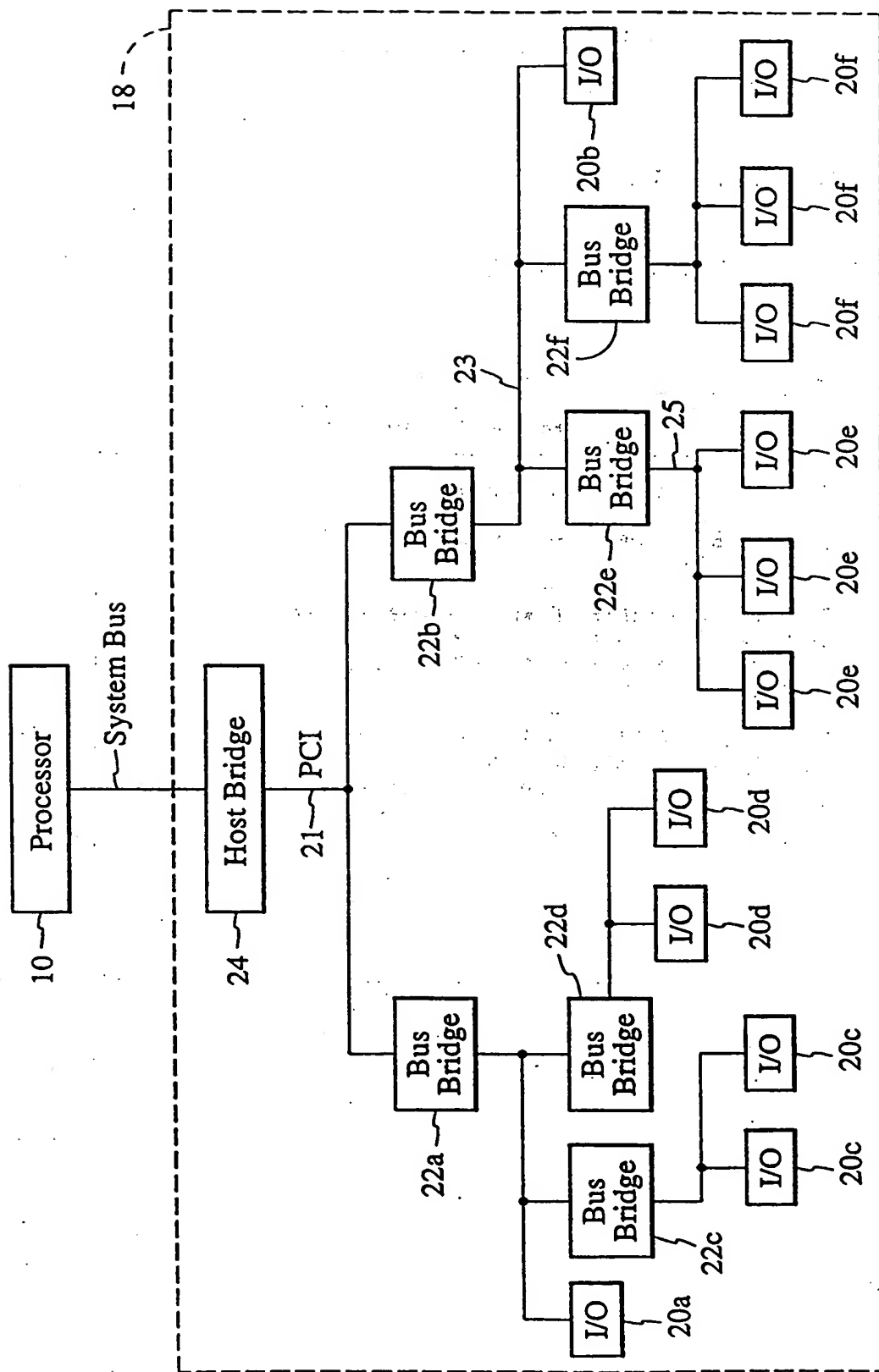


FIG. 2

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/EP 98/01674

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G06F11/22

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	EP 0 811 929 A (COMPAQ COMPUTER CORP) 10 December 1997 see page 57, line 25 - page 59, line 10; figures 1, 26B	1-30
Y	WO 97 00480 A (INTEL CORP) 3 January 1997 see page 17, line 4 - line 19 see page 37, line 1 - page 42, line 1; figures 1, 2	1-30
Y	"BUS FAULT IDENTIFICATION ALGORITHM" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 32, no. 6A, November 1989, pages 153-163, XP000043158 see the whole document	1-30

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

18 August 1998

Date of mailing of the international search report

07/09/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Herreman, G

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 98/01674

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>T. SHANLEY ET AL.: "PCI System Architecture, Third Edition" 1995 , ADDISON-WESLEY PUBLISHING COMPANY , READING, MASSACHUSETTS, US XP002074828 Chapter 10: Error Detection and Handling see whole chapter</p> <p>-----</p>	1-30

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 98/01674

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0811929 A	10-12-1997	JP 10063612 A	06-03-1998
WO 9700480 A	03-01-1997	AU 6334496 A	15-01-1997
		EP 0834135 A	08-04-1998

Form PCT/ISA/210 (patent family annex) (July 1992)